

Article

Not peer-reviewed version

Gumbel–Logistic Unit Distribution with Application in Telecommunications Data Modeling

[Vladica S. Stojanović](#)*, [Mihailo Jovanović](#), [Brankica Pažun](#), Zlatko Langović

Posted Date: 30 August 2024

doi: 10.20944/preprints202408.2209.v1

Keywords: Unit distributions; Gumbel distribution; generalized logistic map; asymmetry; parameter estimation; quantiles; telecommunications




Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Gumbel–Logistic Unit Distribution with Application in Telecommunications Data Modeling

Vladica S. Stojanović ^{1,*} , Mihailo Jovanović ¹, Brankica Pažun ² and Zlatko Langović ³

¹ Department of Informatics & Computer Sciences, University of Criminal Investigation and Police Studies, Belgrade 11000, Serbia

² Department of Informatics, Mathematics and Statistics, Faculty of Engineering Management, Belgrade 11000, Serbia

³ Department of Business Economy, Faculty of Hotel Management and Tourism, University of Kragujevac, Vrnjačka Banja 36210, Serbia

* Correspondence: vladica.stojanovic@kpu.edu.rs

Abstract: The manuscript deals with a new two-parameter unit stochastic distribution, obtained by transforming the Gumbel distribution, using generalized logistic mapping, into a unit interval. The distribution thus obtained is called the Gumbel–Logistic Unit (abbr. GLU) distribution, and its key stochastic properties have been investigated in detail. Among others, it is shown that the GLU distribution, unlike the Gumbel one which is always positively asymmetric, can take both asymmetric forms. Also, the procedure for estimating parameters based on quantiles, along with the asymptotic properties of the obtained estimators and the study of their numerical simulation, is presented. Finally, the application of the GLU distribution in modeling some real–world data related to telecommunications is considered.

Keywords: unit distributions; Gumbel distribution; generalized logistic map; asymmetry; parameter estimation; quantiles; telecommunications

MSC: 60E05; 62E10; 62F10

1. Introduction

One of the most attractive areas of contemporary probability theory are certainly unit stochastic distributions, defined on the interval $(0, 1)$ and commonly used as stochastic models that describe so-called proportional (percentage) variables. Thus, they represent theoretical models that can explain the behavior of some real phenomena (see as some more recent ones, e.g., [1–9]). Still, it is worth pointing out that modeling with unit distributions is very specific, primarily due to the limitation of the data within the $(0, 1)$ interval. Although the procedure for creating unit distributions can be given in a rather general form [10], the most common approach is based on continuous transformations of distributions defined on infinite intervals into a unit interval (see as some recent results, e.g., [11–17]). Motivated with procedures similar as in Stojanović et al. [18,19], here is presented a novel unit distribution, called the Gumbel–logistic unit (GLU) distribution.

This distribution is based on a general logistic mapping of the Gumbel distribution into a unit interval, which gives it flexibility and convenience for describing various kind of empirical distributions. The definition of the GLU distribution as well as its basic stochastic properties, related to its modality, asymmetry, moments, etc. are described in the next Section 2. In addition, the hazard rate and quantile functions of the GLU distribution are also discussed in this section. After that, Section 3 considers the procedure for estimating the parameters of the GLU distribution based on the sample quantiles. The asymptotic properties of thus obtained estimators were also examined, along with an appropriate Monte Carlo numerical study. Section 4 presents the application of the GLU distribution in fitting some real–world data, related to telecommunications and machine learning. Finally, Section 5 provides some concluding highlights.

2. The GLU Distribution

In the first part of this section, the concept of GLU distribution is given as well as some of its basic stochastic properties. Then the hazard rate and quantile function are described, which are used to explore some other properties of the GLU distribution.

2.1. Definition and Key Properties

Let us start from a random variable (RV) Y with a zero-centered Gumbel distribution, whose probability density function (PDF) is:

$$g(y; \beta) = \beta \exp(-\beta y - \exp(-\beta y)), \quad (1)$$

where $y \in \mathbb{R}$ and $\beta > 0$ is the scale parameter. As is known, the Gumbel distribution is a special case of the so-called generalized distribution of extreme values (of the first order) and as such is often used to model the distribution of maxima of a large number of samples from various distributions. Therefore, the basic idea here is to transform this distribution into a unit $(0, 1)$ interval, where a similar application to the newly created distribution can be found. For this purpose, the so-called general logistics map (of the first kind) is used, defined as follows:

$$x = \varphi(y; \alpha) = (1 + \exp(y))^{-1/\alpha},$$

wherein is $\alpha > 0$. Notice that the logistic map is the bijective and continuous function $\varphi : (-\infty, +\infty) \times (0, +\infty) \rightarrow (0, 1)$, with limits:

$$\lim_{y \rightarrow -\infty} \varphi(y; \alpha) = 1^-, \quad \lim_{y \rightarrow +\infty} \varphi(y; \alpha) = 0^+.$$

According to Equation (1) and the inverse function $y = \varphi^{-1}(x; \alpha) = \ln(x^{-\alpha} - 1)$, the novel RV $X = \varphi(Y; \beta)$, defined on the unit interval $(0, 1)$ is obtained. After some computations, the PDF of the RV X can be expressed as follows:

$$f(x; \alpha, \beta) = g(\varphi^{-1}(x; \beta); \alpha) \cdot \left| \frac{\partial \varphi^{-1}(x; \beta)}{\partial x} \right| = \frac{\alpha \beta x^{\alpha\beta-1}}{(1-x^\alpha)^{\beta+1}} \exp\left(-\frac{x^{\alpha\beta}}{(1-x^\alpha)^\beta}\right), \quad (2)$$

where $x \in (0, 1)$ and $\int_0^1 f(x; \alpha, \beta) dx = 1$. Thus, the RV X , with the PDF given by Equation (2), has a Gumbel-logistic unit (GLU) distribution, with the parameters $\alpha, \beta > 0$, which will be further denoted as $X : \mathcal{G}(\alpha, \beta)$. Notice that GLU distribution is a two-parameter distribution, where, in addition to the scale parameter $\beta > 0$, there is also a shape parameter $\alpha > 0$. Also, it is worth noting that the GLU distribution has similar properties to the Gumbel distribution. However, as is known, the Gumbel distribution is unimodal and positively asymmetric, while the GLU distribution has some other special characteristics. We first describe some of them with the following statement.

Theorem 1. Let $X : \mathcal{G}(\alpha, \beta)$ be the RV with the GLU distribution, whose PDF is given by Equation (2). Then, the following statements hold:

(i) When $\alpha\beta > 1$, the RV X is unimodal and both-sides vanishing, that is,

$$\lim_{x \rightarrow 0^+} f(x; \alpha, \beta) = \lim_{x \rightarrow 1^-} f(x; \alpha, \beta) = 0^+. \quad (3)$$

(ii) When $\alpha\beta \leq 1$ and the function

$$P(z; \alpha, \beta) := \alpha \beta z^\beta (z + 1) - \alpha(\beta + 1)z + 1 - \alpha\beta$$

has at least one positive root $z = z_0 > 0$, the RV X is unimodal, left-tailed and right-side vanishing, that is,

$$\lim_{x \rightarrow 0^+} f(x; \alpha, \beta) > 0, \quad \lim_{x \rightarrow 1^-} f(x; \alpha, \beta) = 0^+. \quad (4)$$

(iii) Otherwise, the PDF of the RV X strictly decreases on $(0, 1)$, with boundary properties as in Equation (4).

Proof. We use a procedure similar as in Stojanović et al. [19]. First, after some algebraic computation, the first partial derivative of the function $f(x, \alpha, \beta)$, with respect to $x \in (0, 1)$, is as follows:

$$\frac{\partial f(x; \alpha, \beta)}{\partial x} = \frac{\alpha \beta x^{\alpha \beta - 2}}{(1 - x^\alpha)^{2(\beta + 1)}} \exp\left(-\frac{x^{\alpha \beta}}{(1 - x^\alpha)^\beta}\right) \left((1 - x^\alpha)^\beta (\alpha \beta + (\alpha + 1)x^\alpha - 1) - \alpha \beta x^{\alpha \beta} \right).$$

Thus, it is obvious that the equation $\partial f(x; \alpha, \beta) / \partial x = 0$ is equivalent to $\psi(x; \alpha, \beta) = 0$, where:

$$\psi(x; \alpha, \beta) = (1 - x^\alpha)^\beta (\alpha \beta + (\alpha + 1)x^\alpha - 1) - \alpha \beta x^{\alpha \beta}.$$

After substitution $z = x^\alpha / (1 - x^\alpha) > 0$ and some rearrangement, the previous equation is easily transformed into the equivalent one $P(z; \alpha, \beta) = 0$, i.e.,

$$\alpha \beta z^{\beta + 1} + \alpha \beta z^\beta - \alpha(\beta + 1)z + 1 - \alpha \beta = 0. \quad (5)$$

Based on it, by applying Descartes' rules of signs, the following cases can be observed:

(i) When $\alpha \beta > 1$, there exists exactly one solution $z = z_0$ of Equation (5). Therefore, it obviously represents the unique mode of GLU distribution.

(ii) When $\alpha \beta \leq 1$, Equation (5) can have 0 or 2 different solutions and the other two cases mentioned in the theorem are then obtained.

Finally, according to:

$$\lim_{x \rightarrow 0^+} f(x; \alpha, \beta) = \begin{cases} 0^+, & \alpha \beta > 1 \\ 1, & \alpha \beta = 1 \\ +\infty, & \alpha \beta < 1 \end{cases}, \quad \lim_{x \rightarrow 1^-} f(x; \alpha, \beta) = 0^+,$$

the boundary conditions given by Equations (3) and (4) follows, and the theorem is proved completely. \square

Remark 1. The preceding theorem describes three basic forms of the PDFs of the GLU distribution, which can also be seen in Figure 1(a). As can be easily noticed, the GLU distribution can have various shapes, where in addition to the typical one with a "peak", which is similar to the Gumbel distribution, it can also have a decreasing one. Let us also note that the shape of the GLU distribution depends on the number of solutions to Equation (5). Thus, for the standard Gumbel distribution with $\beta = 1$, this equation becomes quadratic one $\alpha z^2 - \alpha z + 1 - \alpha = 0$, and a mode exists if and only if its discriminant satisfies $\alpha^2 - 4\alpha(1 - \alpha) \geq 0$, i.e., $\alpha \geq 4/5$. Still, a more precise examination of the modality of GLU distribution will be further presented below.

According to Equation (2), the cumulative distribution function (CDF) of the GLU distribution can be obtained as:

$$F(x; \alpha, \beta) := P\{X < x\} = \int_0^x f(t; \alpha, \beta) dt = 1 - \exp\left(-\frac{x^{\alpha \beta}}{(1 - x^\alpha)^\beta}\right), \quad (6)$$

where $0 < x < 1$. Figure 1(b) shows the plots of this function for some parameter values $\alpha, \beta > 0$. Obviously, the function $F(x; \alpha, \beta)$ is well defined at the ends of the unit interval, because is valid:

$$\lim_{x \rightarrow 0^+} F(x; \alpha, \beta) = 0^+, \quad \lim_{x \rightarrow 1^-} F(x; \alpha, \beta) = 1^-,$$

for all $\alpha, \beta > 0$. Using the CDF $F(x; \alpha, \beta)$, one can simply prove the following statement regarding the asymmetry conditions for the GLU distribution:

Theorem 2. Let $X : \mathcal{G}(\alpha, \beta)$ is the RV with the GLU distribution, and

$$\zeta(\beta) := \frac{1}{c} \ln(1 + c^{-1/\beta})$$

the function on $\beta > 0$, where $c := \ln 2 \approx 0.69315$. Then, the GLU distribution is positively asymmetric when $\alpha < \zeta(\beta)$, and vice versa, it is negatively asymmetric when $\alpha > \zeta(\beta)$.

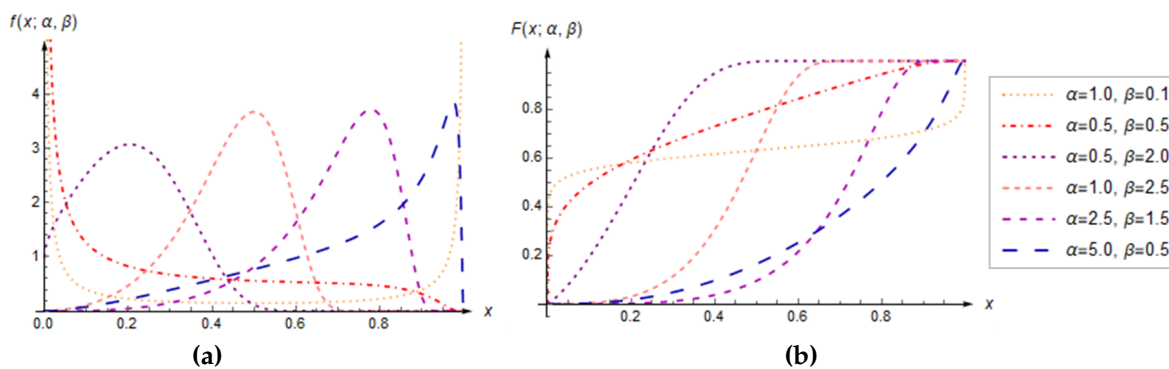


Figure 1. Plots of the PDFs (a) and CDFs (b) of the GLU distribution for different values of parameters $\alpha, \beta > 0$.

Proof. Note that the CDF of the GLU distribution, given by Equation (6), is a strictly increasing function on the interval $(0, 1)$. Therefore, the median $m \in (0, 1)$ can be easily obtained as solutions of the equation:

$$F(m; \alpha, \beta) = \int_0^m f(x; \alpha, \beta) dx = 1 - \exp\left(-\frac{m^{\alpha\beta}}{(1-m^\alpha)^\beta}\right) = \frac{1}{2}. \quad (7)$$

Consequently, the RV $X : \mathcal{G}(\alpha, \beta)$ will be positively asymmetric if and only if is valid $m = F^{-1}(1/2; \alpha, \beta) < 1/2$, that is, $F(1/2; \alpha, \beta) > 1/2$. After some algebraic calculation it is easy to see that the last inequality indeed gives $\alpha < \zeta(\beta)$, and the corresponding inequality for the negatively asymmetric GLU distribution is obtained in a similar way. \square

Remark 2. The previous theorem gives the parameter dependence $\alpha = \zeta(\beta)$ under which the GLU distribution has different skewness. It is shown in Figure 2(a), along with the dependence $\alpha\beta = 1$, mentioned in Theorem 1, which indicating different shapes of the GLU distribution. Thereby, the point $\beta_0 \approx 0.70147$ which represents the solution of the equation $\zeta(\beta) = 1/\beta$ is clearly visible. It should be emphasized that the dependence $\alpha = \zeta(\beta)$ does not indicate the symmetry of the GLU distribution. As an illustration, Figure 2(b) shows graphs of PDFs for which this dependence is applied, that is, for which the median is equal to $m = 1/2$. Note that the "true" symmetry of the GLU distributed RV X does not exist, which also confirms its PDF given by Equation (2). In this way, asymmetry represents one of the essential characteristics of this distribution.

We now describe the procedure for obtaining the r^{th} moments of the GLU distribution:

$$\mu_r(\alpha, \beta) := E(X^r) = \int_0^1 x^r f(x; \alpha, \beta) dx = \int_0^1 x^r dF(x; \alpha, \beta), \quad (8)$$

where $r \in \mathbb{N}$. Note that it is valid:

$$0 < E(X^r) \leq \int_0^1 f(x; \alpha, \beta) dx = 1,$$

so the r^{th} moments exist, but they cannot be expressed in a closed form. To this end, it is convenient to perform the series expansion for $E(X^r)$, given by the following statement.

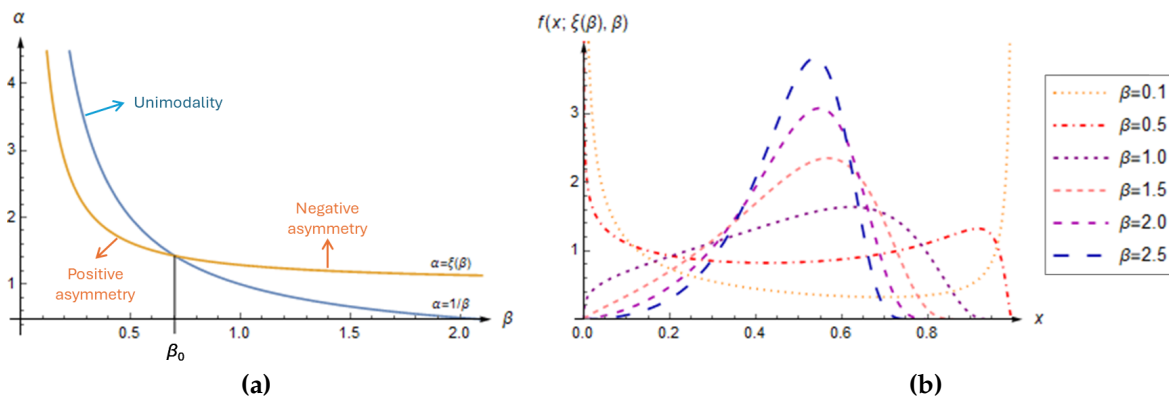


Figure 2. (a) Parameter areas with different shapes and asymmetry of the GLU distribution. (b) Some PDFs of the GLU distributed RV X , where the dependence $\alpha = \zeta(\beta)$ holds.

Theorem 3. The r^{th} moment of the RV X with GLU distribution $\mathcal{G}(\alpha, \beta)$ can be expressed as follows:

$$E(X^r) = \frac{r}{\alpha\beta} \sum_{j=0}^{+\infty} \binom{-r/\alpha - 1}{j} \left[\gamma\left(\frac{r+j\alpha}{\alpha\beta}, 1\right) + \Gamma\left(-\frac{j+1}{\beta}, 1\right) \right], \quad (9)$$

where $\gamma(a, 1) := \int_0^1 t^{a-1} \exp(-t) dt$ and $\Gamma(b, 1) := \int_1^\infty t^{b-1} \exp(-t) dt$ are the lower and upper incomplete gamma functions, respectively.

Proof. According to definition of the r^{th} moment, given by Equation (8), and by applying integration by parts, one obtains:

$$\begin{aligned} E(X^r) &= \int_0^1 x^r dF(x; \alpha, \beta) = \int_1^0 x^r d(1 - F(x; \alpha, \beta)) = r \int_0^1 x^{r-1} (1 - F(x; \alpha, \beta)) dx \\ &= r \int_0^1 x^{r-1} \exp\left[-\left(\frac{x^\alpha}{1-x^\alpha}\right)^\beta\right] dx = \frac{r}{\alpha\beta} \int_0^{+\infty} \frac{t^{\frac{r}{\alpha\beta}-1}}{(1+t^{\frac{1}{\beta}})^{\frac{r}{\alpha}+1}} \exp(-t) dt \\ &= \frac{r}{\alpha\beta} (I_1(r; \alpha, \beta) + I_2(r; \alpha, \beta)), \end{aligned} \quad (10)$$

where $t = (x^\alpha / (1 - x^\alpha))^{1/\beta}$ and

$$I_1(r; \alpha, \beta) = \int_0^1 \frac{t^{\frac{r}{\alpha\beta}-1}}{(1+t^{\frac{1}{\beta}})^{\frac{r}{\alpha}+1}} \exp(-t) dt, \quad I_2(r; \alpha, \beta) = \int_1^{+\infty} \frac{t^{\frac{r}{\alpha\beta}-1}}{(1+t^{\frac{1}{\beta}})^{\frac{r}{\alpha}+1}} \exp(-t) dt. \quad (11)$$

Obviously, both integrals in (11) are absolutely convergent, and applying the generalized binomial formula in them, it follows:

$$\begin{aligned} I_1(r; \alpha, \beta) &= \int_0^1 t^{\frac{r}{\alpha\beta}-1} (1+t^{\frac{1}{\beta}})^{-\frac{r}{\alpha}-1} \exp(-t) dt = \int_0^1 \sum_{j=0}^{+\infty} \binom{-r/\alpha - 1}{j} t^{\frac{r}{\alpha\beta} + \frac{j}{\beta} - 1} \exp(-t) dt, \\ I_2(r; \alpha, \beta) &= \int_1^{+\infty} t^{\frac{r}{\alpha\beta}-1} t^{-\frac{1}{\beta}(\frac{r}{\alpha}+1)} (1+t^{-\frac{1}{\beta}})^{-\frac{r}{\alpha}-1} \exp(-t) dt = \int_1^{+\infty} \sum_{j=0}^{+\infty} \binom{-r/\alpha - 1}{j} t^{-\frac{j+1}{\beta}-1} \exp(-t) dt. \end{aligned} \quad (12)$$

Thereby, the following inequalities hold:

$$\sum_{j=0}^{+\infty} \int_0^1 \left| \binom{-r/\alpha - 1}{j} t^{\frac{r}{\alpha\beta} + \frac{j}{\beta} - 1} \exp(-t) dt \right| \leq \sum_{j=0}^{+\infty} \binom{-r/\alpha - 1}{j} \int_0^1 t^{\frac{r}{\alpha\beta} - 1} dt = \frac{\alpha\beta}{r2^{r/\alpha+1}} < +\infty,$$

$$\sum_{j=0}^{+\infty} \int_1^{+\infty} \left| \binom{-r/\alpha - 1}{j} t^{-\frac{j+1}{\beta} - 1} \exp(-t) dt \right| \leq \sum_{j=0}^{+\infty} \binom{-r/\alpha - 1}{j} \int_1^{+\infty} t^{-\frac{1}{\beta} - 1} dt = \frac{\beta}{2^{r/\alpha+1}} < +\infty,$$

so the order of the sign of integral and sum can be changed (see, e.g., Theorem 1.38 in Rudin [20]). Thus, according to Equations (10) and (12), as well as after some simple calculation, Equation (9) follows. \square

Remark 3. As already mentioned, Equation (9) can be suitable for numerical (approximate) calculation of the moments of GLU distribution, first of all its mean value $E(X) = \mu_1(\alpha, \beta)$ and the variance $\text{Var}(X) = \mu_2(\alpha, \beta) - \mu_1^2(\alpha, \beta)$. For instance, when $\alpha = 2$ and $\beta = 1$ their approximate values (calculated by summing 10^5 terms in Equation (9)) are, respectively, $E(X) \approx 0.605526$ and $\text{Var}(X) \approx 0.404871$. In a similar way one can calculate the other statistical measures of the GLU distribution. Thus, the skewness coefficient and the kurtosis are, respectively,

$$S(\alpha, \beta) = \frac{E(X - \mu_1^3(\alpha, \beta))}{[\text{Var}(X)]^{3/2}} = \frac{\mu_3(\alpha, \beta) - 3\mu_1(\alpha, \beta)\text{Var}(X) - \mu_1^3(\alpha, \beta)}{[\text{Var}(X)]^{3/2}},$$

$$K(\alpha, \beta) = \frac{E(X - \mu_1(\alpha, \beta))^4}{[\text{Var}(X)]^2} = \frac{\mu_4(\alpha, \beta) - 4\mu_3(\alpha, \beta)\mu_1(\alpha, \beta) + 6\mu_2(\alpha, \beta)\mu_1^2(\alpha, \beta) - 3\mu_1^4(\alpha, \beta)}{[\text{Var}(X)]^2}.$$

2.2. Hazard Rate and Quantile Functions

In this part, we consider two important functions which additionally characterize the GLU distribution. First, the hazard rate function (HRF) of the GLU distribution can be easily obtained according to Equations (2) and (6), as follows:

$$H(x; \alpha, \beta) := \frac{f(x; \alpha, \beta)}{1 - F(x; \alpha, \beta)} = \frac{\alpha\beta x^{\alpha\beta-1}}{(1 - x^\alpha)^{\beta+1}}. \quad (13)$$

Obviously, the function $H(x; \alpha, \beta)$ strictly increases when $\alpha\beta \geq 1$, which means that in that case the probability of failure of the designed system increases. On the other hand, when $\alpha\beta < 1$, this HRF has two asymptotic points at $x = 0$ and $x = 1$, i.e., it is a bath-tube shaped. Then, there is a so-called declining failure rate (DFR) that describes a phenomenon where the probability of failure decreases in some interval. Both of these situations can be obtained from the GLU distribution, for certain values of its parameters, as can be also seen in Figure 3(a). Similar to some previous results on unit distributions (see, e.g., Bakouch et al. [4] or Biçer et al. [15]), one typical characterization of the HRF of the GLU distribution can be given by the following statement:

Theorem 4. The RV X has the GLU distribution $\mathcal{G}(\alpha, \beta)$, with the HRF $H(x; \alpha, \beta)$ defined by Equation (13), if and only if this HRF satisfies differential equation:

$$\frac{\partial H(x; \alpha, \beta)}{\partial x} \cdot \frac{1}{(H(x; \alpha, \beta))^2} = \frac{(1 - x^\alpha)^\beta ((\alpha + 1)x^\alpha + \alpha\beta - 1)}{\alpha\beta x^{\alpha\beta}}. \quad (14)$$

Proof. We have use the procedure similar as in Bakouch et al. [4], and first assume that RV X has the GLU distribution $\mathcal{G}(\alpha, \beta)$. Thus, the logarithm of its PDF, given by Equation (2), is as follows:

$$\ln f(x; \alpha, \beta) = \ln(\alpha\beta) + (\alpha\beta - 1) \ln x - (\beta + 1) \ln(1 - x^\alpha) - \frac{x^{\alpha\beta}}{(1 - x^\alpha)^\beta}.$$

By differentiating both sides of the previous equality on x , one obtains:

$$\frac{\partial f(x; \alpha, \beta)}{\partial x} \cdot \frac{1}{f(x; \alpha, \beta)} = \frac{\alpha\beta - 1}{x} + \frac{\alpha(\beta + 1)x^{\alpha-1}}{1 - x^\alpha} - \frac{\alpha\beta x^{\alpha\beta-1}}{(1 - x^\alpha)^{\beta+1}}.$$

Therefore, according to the well-known property of HRF (see, e.g., Salinas et al. [21]), and using Equation (13), it follows:

$$\begin{aligned} \frac{\partial H(x; \alpha, \beta)}{\partial x} \cdot \frac{1}{H(x; \alpha, \beta)} &= \frac{\partial f(x; \alpha, \beta)}{\partial x} \cdot \frac{1}{f(x; \alpha, \beta)} + H(x; \alpha, \beta) \\ &= \frac{\alpha\beta - 1}{x} + \frac{\alpha(\beta + 1)x^{\alpha-1}}{1 - x^\alpha}. \end{aligned} \quad (15)$$

Thus, the last equality and Equation (13), after simple calculation, yield Equation (14).

Otherwise, let us assume that Equation (14) holds, which after integration becomes:

$$-\frac{1}{H(x; \alpha, \beta)} = -\frac{(1 - x^\alpha)^{\beta+1}}{\alpha\beta x^{\alpha\beta-1}},$$

and the HRF $H(x; \alpha, \beta)$ is obtained as in Equation (13). Furthermore, by replacing this function into Equation (15) and after integration, one obtains:

$$\begin{aligned} \ln f(x; \alpha, \beta) &= \int \frac{\partial f(x; \alpha, \beta)}{\partial x} \cdot \frac{dx}{f(x; \alpha, \beta)} = \int \left(\frac{\partial H(x; \alpha, \beta)}{\partial x} \cdot \frac{1}{H(x; \alpha, \beta)} - H(x; \alpha, \beta) \right) dx \\ &= \int \left(\frac{\alpha\beta - 1}{x} + \frac{\alpha(\beta + 1)x^{\alpha-1}}{1 - x^\alpha} - \frac{\alpha\beta x^{\alpha\beta-1}}{(1 - x^\alpha)^{\beta+1}} \right) dx + C_1, \end{aligned}$$

that is

$$f(x; \alpha, \beta) = \frac{\alpha\beta x^{\alpha\beta-1}}{(1 - x^\alpha)^{\beta+1}} \exp\left(C_1 - \frac{x^{\alpha\beta}}{(1 - x^\alpha)^\beta}\right).$$

After another integration, it follows:

$$F(x; \alpha, \beta) = \int f(x; \alpha, \beta) dx + C_2 = C_2 - \exp\left(C_1 - \frac{x^{\alpha\beta}}{(1 - x^\alpha)^\beta}\right),$$

so according to the conditions $F(0; \alpha, \beta) = 0$ and $F(1; \alpha, \beta) = 1$ it is easy to obtain that $C_1 = 0$ and $C_2 = 1$. Thus, the function $F(x; \alpha, \beta)$ is indeed the CDF of RV $X : \mathcal{G}(\alpha, \beta)$, and theorem is completely proved. \square

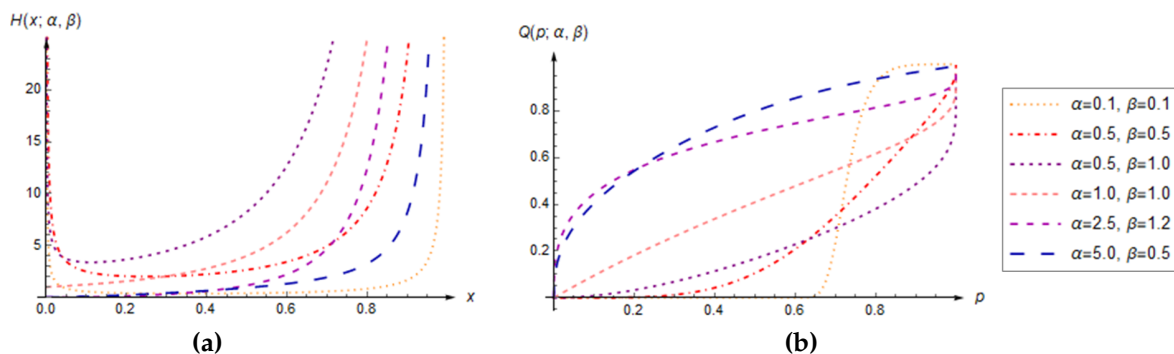


Figure 3. Plots of the HRF (a) and QF (b) of the GLU distribution, obtained for some parameters values $\alpha, \beta > 0$.

In the following, the so-called quantile function (QF) of the GLU distribution is considered, obtained as inverse function of its CDF:

$$Q(p; \alpha, \beta) := F^{-1}(p; \alpha, \beta) = \left(1 - \frac{1}{1 + (-\ln(1-p))^{1/\beta}}\right)^{1/\alpha}. \quad (16)$$

Herein is $p = F(x; \alpha, \beta)$ and plots of the QF for some parameter values are shown in Figure 3(b). The QF provides a useful tool for further investigating the properties of the GLU distribution, as given in the following statement.

Theorem 5. Let $X : \mathcal{G}(\alpha, \beta)$ is the GLU distributed RV, whose QF is defined by Equation (16). Then, the RV X is unimodal if and only if its parameters (α, β) satisfy the equality:

$$\alpha = \left[\beta(1 + \ln(1-p_0)) + (-\ln(1-p_0))^{1/\beta}(1 + \beta + \beta \ln(1-p_0)) \right]^{-1}, \quad (17)$$

where $\partial^3 Q(p_0; (\alpha, \beta)) / \partial p^3 > 0$, $p_0 = F(x_0; \alpha, \beta)$ and $x_0 \in (0, 1)$ is the mode of the GLU distribution. Otherwise, the PDF of X is a strictly decreasing function.

Proof. Applying the derivative rule of the inverse function, the derivatives of QF $Q(p; \alpha, \beta)$ up to the third order are as follows:

$$\begin{aligned} \frac{\partial Q(p; \alpha, \beta)}{\partial p} &= \frac{1}{\partial F(x; \alpha, \beta) / \partial x} = \frac{1}{f(x; \alpha, \beta)}, \\ \frac{\partial^2 Q(p; \alpha, \beta)}{\partial p^2} &= -\frac{\partial f(x; \alpha, \beta) / \partial x}{f^2(x; \alpha, \beta)}, \\ \frac{\partial^3 Q(p; \alpha, \beta)}{\partial p^3} &= \frac{2(\partial f(x; \alpha, \beta) / \partial x)^2 - f(x; \alpha, \beta) \partial^2 f(x; \alpha, \beta) / \partial x^2}{f^3(x; \alpha, \beta)}, \end{aligned}$$

where $x = Q(p; \alpha, \beta)$. At the same time, if the value $x = x_0$ is the mode for the RV X , it is the critical point of its PDF $f(x; \alpha, \beta)$, that is, the value $p_0 = F(x_0; \alpha, \beta)$ satisfies:

$$\frac{\partial^2 Q(p_0; \alpha, \beta)}{\partial p^2} = 0, \quad \frac{\partial^3 Q(p_0; \alpha, \beta)}{\partial p^3} = -\frac{\partial^2 f(x_0; \alpha, \beta) / \partial x^2}{f^2(x_0; \alpha, \beta)} > 0. \quad (18)$$

After some computation, it is obtained:

$$\frac{\partial^2 Q(p; \alpha, \beta)}{\partial p^2} = \frac{(-\ln(1-p))^{\frac{1}{\alpha\beta}-2} (\alpha\beta + \alpha\beta \ln(1-p) + \alpha(-\ln(1-p))^{1/\beta}(1 + \beta + \beta \ln(1-p)) - 1)}{\alpha^2 \beta^2 (1-p)^2 (1 + (-\ln(1-p))^{1/\beta})^{\frac{1}{\alpha}+2}}, \quad (19)$$

so that Equations (18) and (19) obviously give the statement of the theorem. \square

Remark 4. Note that Theorems 1 and 5 provide complete insight into the modality property of the GLU distribution. It is explicitly described by the functional dependence in Equation (17), which is close to logarithmic spirals, as shown in Figure 4. Here are presented, with two different "views", polar plots of parameter dependences (α, β) that give unimodal GLU distribution, for different but fixed values $p_0 \in (0, 1)$. In addition, the QF can also describe the asymmetry conditions of the GLU distribution. The positive and negative asymmetry conditions are easily obtained by solving the inequalities $Q(1/2; \alpha, \beta) < 1/2$ and $Q(1/2; \alpha, \beta) > 1/2$, respectively, and it can simply be proved that they are the same as in Theorem 2.

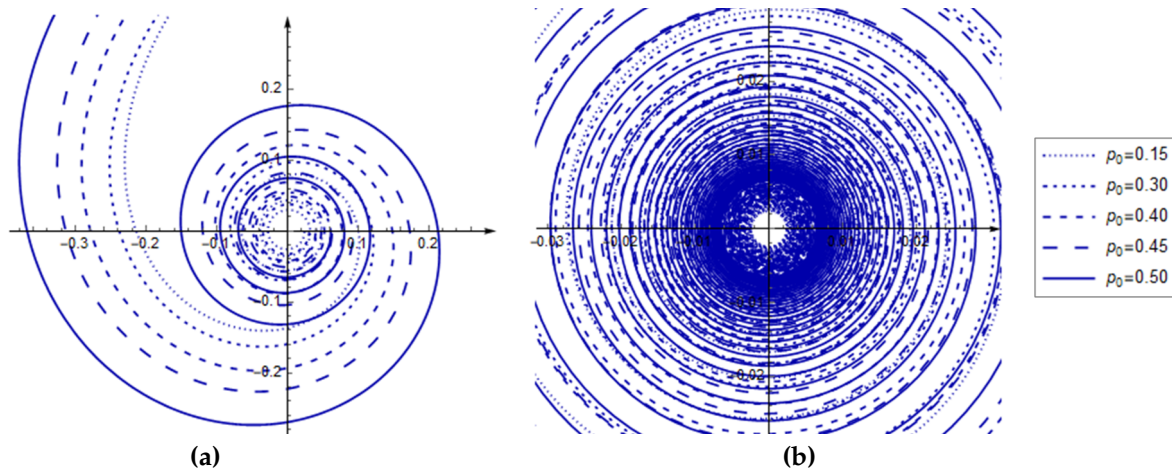


Figure 4. Polar plots of parameter dependences yielding a unimodal GLU distribution, with some fixed values of $p_0 \in (0, 1)$ and two different angular intervals: **(a)** $\beta \in [0, 8\pi]$; **(b)** $\beta \in [0, 80\pi]$.

3. Parameters Estimation & Numerical Simulation Study

In this Section the estimation of the unknown parameters $\alpha, \beta > 0$ of the GLU-distributed RV X is described, based on its observed random sample x_1, \dots, x_n of length n . Let us note first that according to the aforementioned properties of the GLU distribution, some common procedures for estimating its parameters are not appropriate here. For instance, according to Equation (9), it follows that the moments of the GLU distribution cannot be expressed in a closed form, and thus the method of moments cannot be successfully applied. Similarly, the maximum likelihood (ML) estimation method is also associated with certain difficulties, related to numerically finding solutions (α, β) that maximize the likelihood function:

$$L(\alpha, \beta | x_1, \dots, x_n) = \prod_{i=1}^n f(x_i; \alpha, \beta) = (\alpha\beta)^n \prod_{i=1}^n \left[\frac{x_i^{\alpha\beta-1}}{(1-x_i^\alpha)^{\beta+1}} \exp\left(-\frac{x_i^{\alpha\beta}}{(1-x_i^\alpha)^\beta}\right) \right].$$

For these reasons, and similar as in Stojanović et al. [18,19], here we consider parameter estimation methods based on the quantiles of the GLU distribution. These estimators (we will call them Q-estimators) are explicitly given and also have some convenient asymptotic properties, which will be shown below.

In that aim, for a random sample X_1, X_2, \dots, X_n of the length n , let us define the appropriate order statistics $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$. Then the PDF of the i -th order statistic $X_{(i)}$, as is known, can be expressed as follows:

$$f_{X_{(i)}}(x; \alpha, \beta) = \frac{n!}{(i-1)!(n-i)!} f(x; \alpha, \beta) [F(x; \alpha, \beta)]^{i-1} [1 - F(x; \alpha, \beta)]^{n-i}, \quad (20)$$

where $i = 1, \dots, n$. On the other hand, by replacing $p = p_0$ in the QF $Q_p := K(p; \alpha, \beta)$, given by Equation (16), the quantile Q_{p_0} is obtained. Therefore, the appropriate sample quantile can be obtained according to the equality:

$$\hat{Q}_{p_0} := X_{(i_{p_0})}, \quad i_{p_0} = \begin{cases} np_0, & np_0 = 1, 2, \dots, n, \\ 1 + [np_0], & \text{otherwise,} \end{cases} \quad (21)$$

where $[np_0]$ is the integer part of np_0 . Thus, sample quantiles are actually the order statistics, so their distribution is determined by Equation (20).

In order to determine the Q-estimators of the parameters of the GLU-distributed RV X , notice first that for $p_0 = 1 - e^{-1}$ is obtained the quantile $Q_{p_0} = 2^{-1/\alpha}$. Hence, by equating this quantile with sample one \widehat{Q}_{p_0} , the estimator of the shape parameter α is simply obtained as follows:

$$\widehat{\alpha} = -\frac{\ln 2}{\ln \widehat{Q}_{p_0}}. \quad (22)$$

Furthermore, by substituting $p = 1/2$ into the QF $Q(p; \alpha, \beta)$, it is obtained the median of the GLU distribution:

$$m = Q(1/2; \alpha, \beta) = \left(1 - \frac{1}{(\ln 2)^{1/\beta} + 1}\right)^{1/\alpha}. \quad (23)$$

Thus, by equating median with the sample one $\widehat{m} = \widehat{Q}_{1/2}$, and using the estimator $\widehat{\alpha}$, for the estimator of the scale parameter β one obtains:

$$\widehat{\beta} = \frac{\ln(\ln 2)}{\ln(\widehat{m}^{\widehat{\alpha}} / (1 - \widehat{m}^{\widehat{\alpha}}))}. \quad (24)$$

In the following, some asymptotic properties of the proposed estimators are examined:

Theorem 6. *Statistics $(\widehat{\alpha}, \widehat{\beta})$ are consistent and asymptotic normal (AN) estimators of the true parameters (α, β) .*

Proof. To prove the consistency of the proposed estimators, we apply some general results of sample quantile theory. Let us first note that the CDF $p = F(x; \alpha, \beta)$ is a differentiable and increasing function on $p \in (0, 1)$. Therefore, the quantiles $Q_p := Q(p; \alpha, \beta)$ are uniquely determined by Equation (16), while the sample quantiles \widehat{Q}_p are uniquely determined by Equation (21). Now, according to Bahadur's representation of sample quantiles (see, e.g., Theorem 1 in [22], or Serfling [23], pp. 91-92), it follows:

$$\widehat{Q}_p = Q_p + \frac{p - F_n(Q_p)}{f(Q_p; \alpha, \beta)} + \mathcal{O}\left[\left(n^{-1} \ln n\right)^{3/4}\right], \quad (25)$$

where $F_n(x) := n^{-1} \sum_{i=1}^n 1(X_i < x)$ is the empirical CDF of the GLU-distributed RV X . It is well known that for arbitrary $x \in \mathbb{R}$, the empirical CDF $F_n(x)$ almost surely and uniformly converges to the CDF $p = F(x; \alpha, \beta)$, when $n \rightarrow \infty$. Applying this convergence on Equation (25), when $x = Q_p$, one obtains:

$$\widehat{Q}_p \xrightarrow{as} Q_p, \quad n \rightarrow \infty,$$

i.e., the sample quantiles are consistent estimators of the theoretical ones. At the same time, the estimators $(\widehat{\alpha}, \widehat{\beta})$ are continuous functions of the sample quantile \widehat{Q}_{p_0} , where $p_0 = 1 - e^{-1}$, as well as the sample median $\widehat{m} = \widehat{Q}_{1/2}$. Thus, applying the continuity property of almost sure convergence (see, e.g., Serfling [23], p. 24), it follows:

$$(\widehat{\alpha}, \widehat{\beta}) \xrightarrow{as} (\alpha, \beta), \quad n \rightarrow \infty,$$

i.e., $(\widehat{\alpha}, \widehat{\beta})$ are indeed consistent estimators of (α, β) .

We now prove the AN properties of the proposed estimators. To this end, note that under the above assumptions, Equation (25) implies the following convergence in the distribution:

$$\sqrt{n}(\widehat{Q}_p - Q_p) \xrightarrow{d} \mathcal{N}\left(0, \frac{p(1-p)}{[f(Q_p, \alpha, \beta)]^2}\right), \quad n \rightarrow \infty. \quad (26)$$

By using Equation (26), for the sample quantile \widehat{Q}_{p_0} , where $p_0 = 1 - e^{-1}$, one obtains:

$$\sqrt{n}(\widehat{Q}_{p_0} - Q_{p_0}) \xrightarrow{d} \mathcal{N}(0, \sigma_{p_0}^2), \quad n \rightarrow \infty,$$

where, according to Equation (2) and after some calculations, we get:

$$\sigma_{p_0}^2 := \text{Var}(\widehat{Q}_{p_0}) = \frac{p_0(1-p_0)}{[f(p_0; \alpha, \beta)]^2} \Big|_{p_0=1-e^{-1}} = \frac{e-1}{\alpha^2 \beta^2 4^{\frac{1}{\alpha}+1}}.$$

Hence, applying the continuity of convergence in the distribution (see, e.g., Serfling [23], p. 118), for the estimator $\widehat{\alpha}$, defined by Equation (22), it is obtained:

$$\sqrt{n}(\widehat{\alpha} - \alpha) \xrightarrow{d} \mathcal{N}(0, \sigma_{\alpha}^2), \quad n \rightarrow \infty, \quad (27)$$

where:

$$\sigma_{\alpha}^2 := \text{Var}(\widehat{\alpha}) = \left[-\frac{\partial}{\partial Q_{p_0}} \left(\frac{\ln 2}{\ln Q_{p_0}} \right) \right]^2 \Big|_{Q_{p_0}=2^{-1/\alpha}} \cdot \sigma_{p_0}^2 = \frac{\ln^2 2}{Q_{p_0}^2 \ln^4 Q_{p_0}} \Big|_{Q_{p_0}=2^{-1/\alpha}} \cdot \frac{e-1}{\alpha^2 \beta^2 4^{\frac{1}{\alpha}+1}} = \frac{\alpha^2(e-1)}{4 \beta^2 \ln^2 2}.$$

In a similar way, the AN property of the estimator $\widehat{\beta}$, given by Equation (24), is proved. To that end, let us first notice that Equation (26), applying on the sample median $\widehat{m} = \widehat{Q}_{1/2}$, gives the following convergence:

$$\sqrt{n}(\widehat{m} - m) \xrightarrow{d} \mathcal{N}(0, \sigma_m^2), \quad n \rightarrow \infty.$$

Here, according to Equations (23) and (26), as well as after some computations, it follows:

$$\sigma_m^2 := \text{Var}(\widehat{m}) = \frac{1}{4[f(m; \alpha, \beta)]^2} = \frac{(\ln 2)^{\frac{2}{\alpha\beta}-2}}{\alpha^2 \beta^2 (1 + (\ln 2)^{\frac{1}{\beta}})^{\frac{2}{\alpha}+2}}. \quad (28)$$

By applying again the continuity of convergence in distribution, one obtains:

$$\sqrt{n}(\widehat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(0, \sigma_{\beta}^2), \quad n \rightarrow \infty, \quad (29)$$

where, according to Equations (23), (24) and (28), it follows:

$$\begin{aligned} \sigma_{\beta}^2 := \text{Var}(\widehat{\beta}) &= \left[-\frac{\partial}{\partial m} \left(\frac{\ln(\ln 2)}{\ln(m^{\alpha}/(1-m^{\alpha}))} \right) \right]^2 \cdot \sigma_m^2 = \frac{\ln^2(\ln 2)}{m^{2\alpha}(1-m^{\alpha})^2 \ln^4(m^{\alpha}/(1-m^{\alpha}))} \cdot \frac{(\ln 2)^{\frac{2}{\alpha\beta}-2}}{\alpha^2 \beta^2 (1 + (\ln 2)^{\frac{1}{\beta}})^{\frac{2}{\alpha}+2}} \\ &= \frac{\beta^4 (\ln 2)^{-\frac{2}{\beta}} (1 + (\ln 2)^{\frac{1}{\beta}})^4 (\ln 2)^{\frac{2}{\alpha\beta}-2}}{\alpha^2 \beta^2 \ln^2(\ln 2) (1 + (\ln 2)^{\frac{1}{\beta}})^{\frac{2}{\alpha}+2}} = \frac{\beta^2 (1 + (\ln 2)^{1/\beta})^{2-\frac{2}{\alpha}}}{\alpha^2 \ln^2(\ln 2) (\ln 2)^{2+\frac{2}{\beta}-\frac{2}{\alpha\beta}}}. \end{aligned}$$

In this way, the convergences proved in Equations (27) and (29) confirm the AN properties of both estimators $\widehat{\alpha}$ and $\widehat{\beta}$. \square

In the following, a numerical study examining the efficiency of the proposed Q-estimators is presented, based on independent Monte Carlo simulations of samples x_1, \dots, x_n drawn from the GLU distribution. In other words, various samples and parameter values from the GLU distribution were considered, according to which the Q-estimators were calculated, and their statistical analysis was also performed. To that aim, three different samples from the GLU-distribution are examined (also shown in Figure 5, below):

(i) Sample I is taken from a decreasing GLU distribution, with parameters $\alpha = 0.5$ and $\beta = 1$, which satisfies the inequality $\alpha\beta < 1$.

(ii) Sample II is taken from a unimodal, positively skewed GLU distribution, with parameters $\alpha = 0.5$ and $\beta = 2$, so the equality $\alpha\beta = 1$ holds.

(iii) Sample III is taken from a unimodal, negatively skewed GLU distribution, with parameters $\alpha = 2$ and $\beta = 1.5$, which satisfies the inequality $\alpha\beta > 1$.

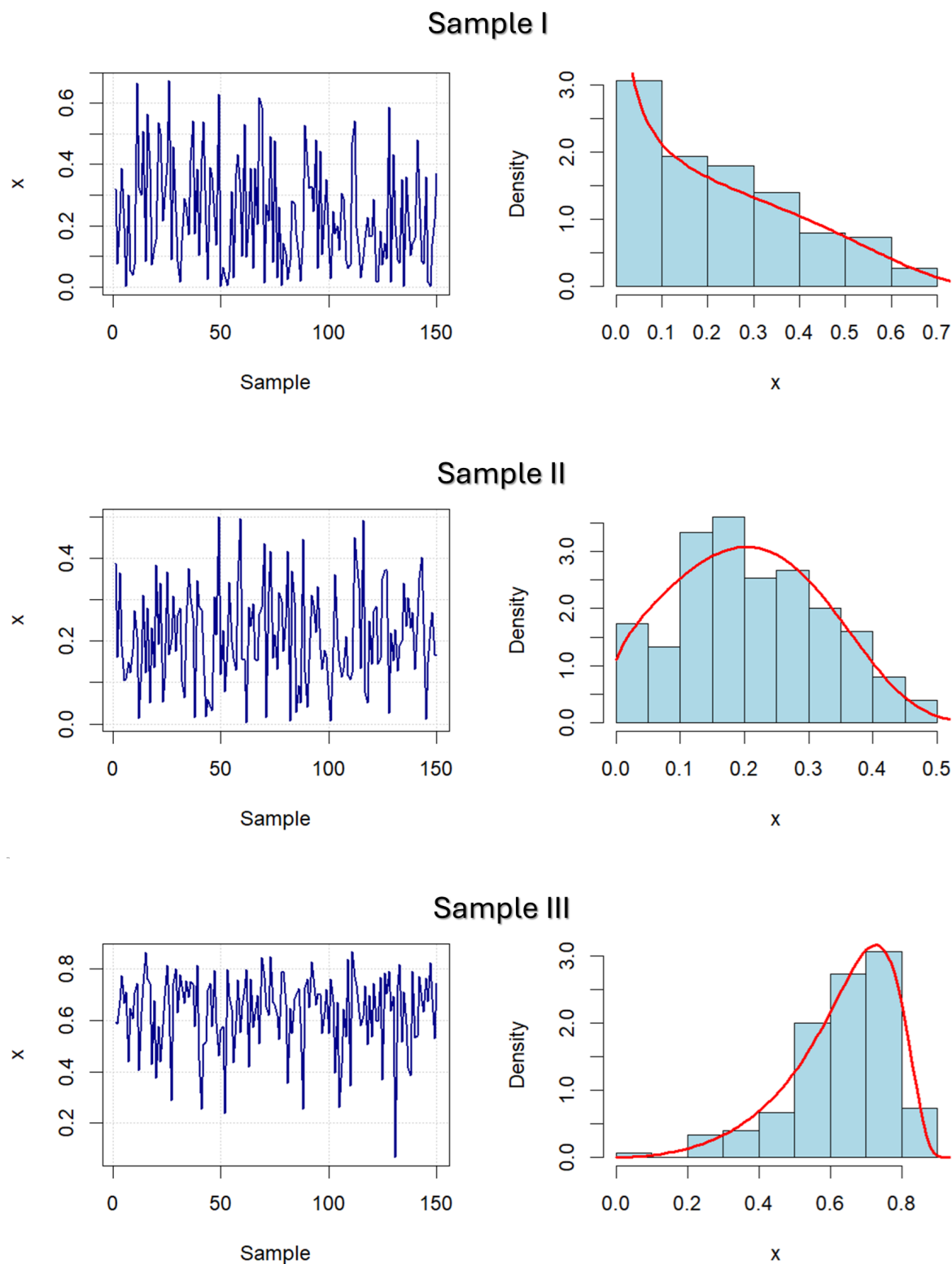


Figure 5. Left plots: Observations of various samples drawn from the GLU-distribution. Right plots: Empirical and fitted PDFs of the RV $X : \mathcal{G}(\alpha, \beta)$.

Note that the simulated sample values are generated by the R-package "distr" [24], and thereafter the Q-estimates $\hat{\alpha}$ and $\hat{\beta}$ are calculated using the procedure described above. To additionally check the efficiency of proposed estimators, realizations of samples of different lengths $n \in \{150, 500, 1500\}$ were considered, so that they are close to the lengths of some of the real-world data that will be analyzed below. In addition, for each of the samples, $S = 250$ independent simulations were conducted, on which an appropriate statistical analysis of the obtained estimates was then performed. The results of this analysis are presented in the following Tables 1–3.

Table 1. Summary statistics, estimation errors, and AN testing of parameter estimates of the GLU distribution: Sample I with the true parameter values $\alpha = 0.5$ and $\beta = 1$.

Statistics	$n = 150$		$n = 500$		$n = 1500$	
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
Min.	0.4015	0.5716	0.4237	0.6739	0.4660	0.8195
Mean	0.5122	1.1906	0.5111	1.1519	0.5020	1.0488
Max.	0.6101	1.8720	0.5807	1.7959	0.5515	1.2252
SD	0.0388	0.9769	0.0221	0.2341	0.0122	0.0439
MSEE	0.0403	0.3218	0.0252	0.1906	0.0168	0.0519
FEE (%)	8.0623	32.177	5.0333	19.062	3.3707	5.1930
AD	0.2933	0.5094	0.3035	0.6068	0.2930	0.4860
(<i>p</i> -value)	(0.5997)	(0.1962)	(0.5704)	(0.1137)	(0.6004)	(0.2240)
W	0.9929	0.9876 *	0.9946	0.9884 *	0.9957	0.9902
(<i>p</i> -value)	(0.2815)	(0.0299)	(0.5138)	(0.0414)	(0.7109)	(0.0890)

* $0.01 < p < 0.05$.

Table 2. Summary statistics, estimation errors, and AN testing of parameters estimates of the GLU distribution: Sample II with the true parameter values $\alpha = 0.5$ and $\beta = 2$.

Statistics	$n = 150$		$n = 500$		$n = 1500$	
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
Min.	0.4299	1.2050	0.4731	1.3330	0.4821	1.5590
Mean	0.4993	2.1085	0.4995	2.0420	0.5000	2.0206
Max.	0.5615	2.6790	0.5194	2.2570	0.5155	2.1400
SD	0.0212	1.0480	9.26×10^{-3}	0.3665	5.58×10^{-3}	0.2013
MSEE	0.0181	0.3746	0.0105	0.0862	6.22×10^{-3}	0.0356
FEE (%)	3.6132	18.728	2.0931	4.3068	1.2481	1.7883
AD	0.3802	1.0201 *	0.2678	0.4337	0.2090	0.3384
(<i>p</i> -value)	(0.4005)	(0.0108)	(0.6826)	(0.2998)	(0.8621)	(0.5021)
W	0.9914	0.9888 *	0.9939	0.9900	0.9956	0.99031
(<i>p</i> -value)	(0.1504)	(0.0489)	(0.4049)	(0.0834)	(0.7032)	(0.0949)

* $0.01 < p < 0.05$.

More specifically, the above Tables contain summary statistics of the calculated estimates, that is, their minimums (Min.), mean values (Mean) and maximums (Max.). In addition, some error statistics are also shown, i.e., the standard deviations (SD), the mean square errors of estimation (MSEE) and fractional errors of estimation (FEE). Finally, the results of the Anderson-Darling and Shapiro-Wilk normality tests are also given. Based on the results obtained in this way, it can be noted that the proposed estimators are efficient, because the bias, the sample range (Max.–Min.), as well as the values of SD, MSEE and FEE decrease with the increase in the sample size. At the same time, it can be noted that stability and efficiency are more significant at the estimates $\hat{\alpha}$, especially in the first sample. This obviously follows from the fact that the estimate $\hat{\beta}$ is calculated by a two-stage procedure, i.e., by using the previously obtained estimate $\hat{\alpha}$.

Table 3. Summary statistics, estimation errors, and AN testing of parameters estimates of the GLU distribution: Sample III with the parameter values $\alpha = 2$ and $\beta = 1.5$.

Statistics	$n = 150$		$n = 500$		$n = 1500$	
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
Min.	1.6710	1.0773	1.8330	1.1057	1.8621	1.1950
Mean	1.9952	1.4901	1.9978	1.5052	2.0010	1.4970
Max.	2.2941	1.9122	2.1605	1.7935	2.1072	1.7245
SD	0.0949	0.6647	0.0606	0.2910	0.0326	0.1534
MSEE	0.0949	0.1901	0.0523	0.0796	0.0327	0.0482
FEE (%)	4.7450	12.675	2.6488	5.3096	1.6375	3.2158
AD	0.3238	2.0687 **	0.2153	0.5059	0.3041	0.5160
(p -value)	(0.5235)	(2.83×10^{-5})	(0.8460)	(0.2001)	(0.5686)	(0.1889)
W	0.99376	0.9806 **	0.9950	0.9903	0.9952	0.9908
(p -value)	(0.3865)	(1.74×10^{-3})	(0.5840)	(0.0949)	(0.6194)	(0.1189)

** $p < 0.01$.

Similar conclusions can be made according to the results of AN testing of these estimates. As previously mentioned, AN testing was conducted using Anderson–Darling and Shapiro–Wilk normality tests, whose statistics, labelled by AD and W respectively, as well as the appropriate p -values were calculated using the R-package "nortest" [25]. According to the results obtained in this way, also presented in Tables 1–3, it can be noted that estimates $\hat{\alpha}$ have a pronounced AN feature, which applies to all observed samples. On the other hand, estimates $\hat{\beta}$ have a less pronounced AN feature, primarily in the case of smaller samples. Nevertheless, the AN property is clearly confirmed in most cases with them as well. Some confirmations of these facts can be also seen in Figure 5, where the realizations of the samples, as well as their empirical and theoretical distributions, are shown.

4. Applications of the GLU Distribution

As mentioned earlier, the Gumbel distribution can be used to model the extremes of some sample values. In more detail, in his original work Gumbel [26] has proved that the maximum of a sample taken from a population with an exponential distribution, after a simple transformation, approaches the Gumbel distribution with increasing sample size. This procedure can also be applied in some practical cases, such as in Burke et al. [27], where the Gumbel distribution is used to analyze maximum rainfall. In a similar way, the maximum load on the telecommunications system, which allows administrators to optimize network capacity and minimize the occurrence of overloads can be modeled with the Gumbel distribution. It is worth pointing that some other application can be in risk analysis related to the ICT technology, which enables companies to better understand and prepare for extreme scenarios that can significantly affect the business by developing methods for the development and introduction of new technologies and concepts, as is shown in Pažun & Langović [28].

For those reasons, this section considers some possible applications of the GLU distribution in real-world data modeling, primarily in the domains of telecommunications and machine learning. The datasets observed here were downloaded from the website "Kaggle.com" [29], a platform focused on analysing and sharing datasets related to machine learning and online data science. Thereby, the observed data represent parts of the training data related to network and telecommunication traffic in India, i.e., describe the satisfaction and participation of end users, as well as the adaptability and extensibility of the corresponding network transport. More specifically, the three real-world datasets analyzed below can be briefly described as follows:

(i) The first data set, named Series A, consists of $n = 152$ data representing the percentage of service usage time of end users. The data was collected by Mirza [30] and as already mentioned, it is a part of the training data intended for machine learning and online coding and modeling.

(ii) The second one (Series B), are also part of the same training data as above, and consists of $n = 542$ monthly end user fees (in Indian Rupees). In doing so, these data are normalized in relation to their maximum and minimum values, and in this way a set of data in a unit interval is obtained.

(iii) Finally, the third set of data, designated as Series C, is obtained from training data authored by Mnassri [31], intended for the development of appropriate predictive models, i.e., training, cross-validation and performance testing of machine learning models. Therefore, Series C consists of $n = 1564$ data, which represent the total daily call length of end users (expressed in minutes), whereby the normalized values are obtained as the ratio of the call duration to the maximum call length.

Realizations of these series are shown in Figure 6(a), while Figure 6(b) shows the values of their corresponding autocorrelation functions (ACFs). As can be easily seen, the ACF values of all series are relatively low, so they can be considered as independent realizations of some unit RV, that is, they can obviously be modeled by one of the unit stochastic distributions.

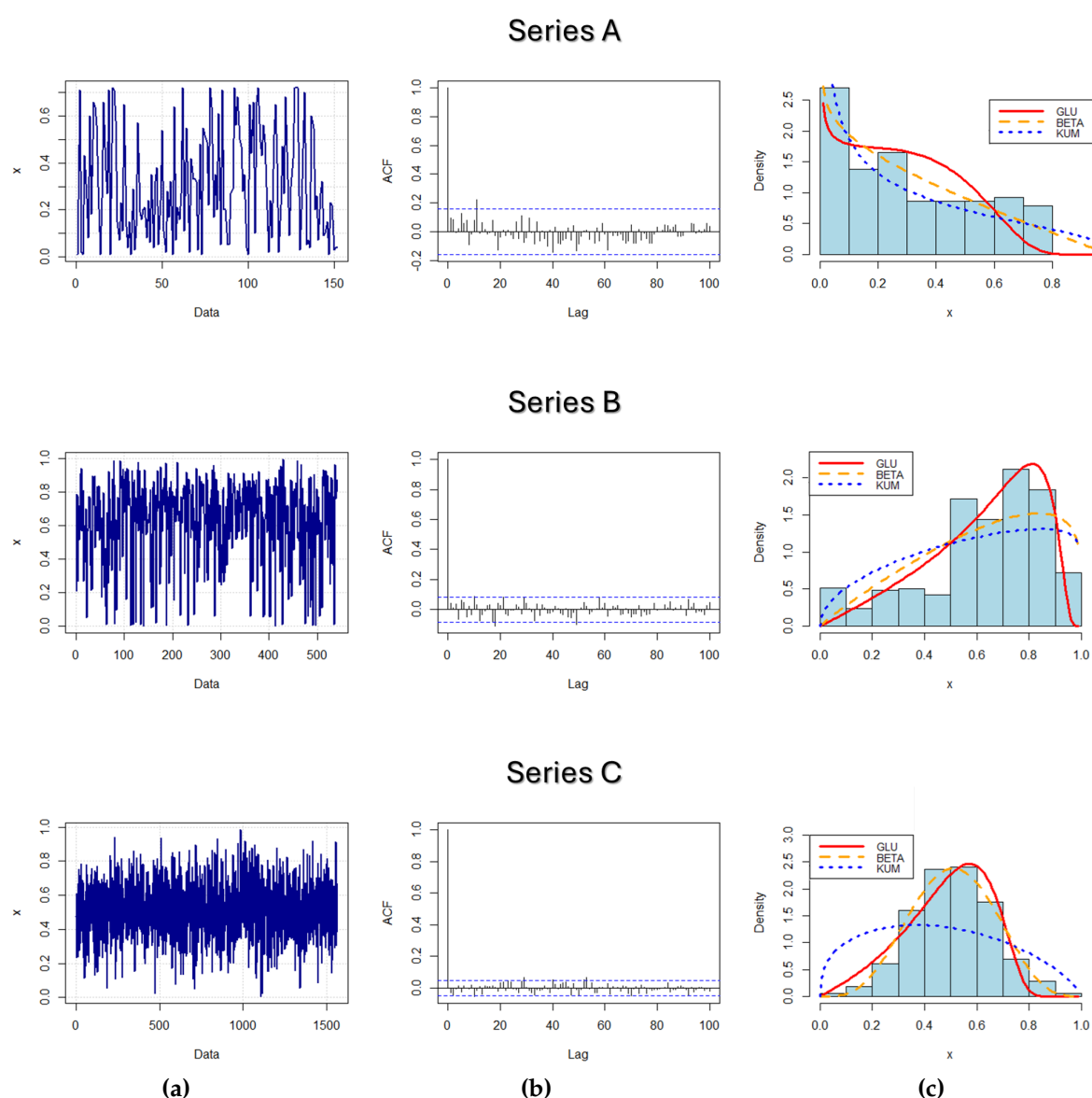


Figure 6. (a): Observed sample values of three real-world data. (b): Estimated ACFs of observed samples (data series). (c): Empirical and fitted PDFs, obtained using the GLU, Beta and Kumaraswamy distributions.

In order to additionally verify the effectiveness of such modeling, in addition to the proposed GLU distribution, some other existing, well-known unit distributions were used to fit the empirical distributions of observed data. In more detail, the GLU distribution is compared with two well-known unit distributions, namely the Beta and Kumaraswamy distributions. Their PDFs are, respectively,

$$g_1(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad g_2(x; \alpha, \beta) = \alpha \beta x^{\alpha-1} (1-x^\alpha)^{\beta-1},$$

where $0 < x < 1$, as $\alpha, \beta > 0$ are distribution parameters, and $B(\alpha, \beta)$ is the beta function. To obtain the estimated parameter values of the Beta distribution, the method of moments (MM) is used here (see, e.g., Kachiashvili & Melikdzhanjan [32]). According to this, the MM estimates are as follows:

$$\hat{\alpha} = \bar{x} \left(\frac{\bar{x}(1-\bar{x})}{\bar{v}} - 1 \right), \quad \hat{\beta} = (1-\bar{x}) \left(\frac{\bar{x}(1-\bar{x})}{\bar{v}} - 1 \right),$$

where $\bar{x} = n^{-1} \sum_{i=1}^n x_i$ and $\bar{v} = (n-1)^{-1} \sum_{i=1}^n (x_i - \bar{x})^2$ are the sample mean and variance, respectively, and the inequality $\bar{v} < \bar{x}(1-\bar{x})$ holds. On the other hand, the maximum likelihood (ML) estimation method is used for the Kumaraswamy distribution, according to which estimators are obtained as solutions of coupled equations (see, e.g., Dey et al. [33]):

$$\frac{n}{\alpha} + \sum_{i=1}^n \ln x_i - (\beta-1) \sum_{i=1}^n \frac{x_i^\alpha \ln x_i}{1-x_i^\alpha} = 0,$$

$$\frac{n}{\beta} + \sum_{i=1}^n \ln(1-x_i^\alpha) = 0.$$

As is known, the proposed estimators for both of the above distributions have the properties of stability and asymptotic normality. In this way, one of the reasons for choosing them is the comparison not only with respect to their distribution, but also with respect to different estimation procedures.

The results of the previously described estimation procedures can be seen in Figure 6(c), where the empirical distributions of the observed data (given by histograms) are shown, along with the corresponding fitted PDFs. As can be easily seen, the empirical distribution of Series A is significantly positively skewed and it is fitted with decreasing theoretical PDFs. On the other side, Series B and Series C have negatively skewed unimodal distributions, with the distribution of Series C being "approximately symmetric" to some extent. This can also be confirmed by the estimated parameter values for each series and for all competing models, which are shown in Table 4 below.

Based on them, $S = 1000$ independent Monte Carlo simulations of the corresponding theoretical distributions were performed and the agreement between the empirical and fitted distributions was checked in several different ways. Namely, the mean square estimation error (MSEE) statistics, the Akaike information criterion (AIC), as well as the Bayesian information criterion (BIC) for model selection were used for this purpose. In addition, the Kolmogorov–Smirnov (KS) test of equivalence of the asymptotic distribution of the two samples was also performed, and all these values are also shown in Table 4.

According to the results obtained in this way, it is noticeable, for instance, that in the case of Series A all three theoretical distributions can be adequate for fitting. On the contrary, for the other two series, the values of MSEE, AIC and BIC are generally lower in cases where GLU and Beta distribution are applied as appropriate fitting models. Nevertheless, it is clear that the GLU distribution has better fitting characteristics than both other theoretical distributions, even in the case of Series C, which has an "approximately symmetric" distribution. Moreover, only with the GLU distribution, the KS test statistics do not reject, with a significant level $p > 0.01$, the hypothesis of the equivalence of the theoretical and the observed empirical distribution.

Table 4. Estimated parameters of the LLU, beta, and Kumaraswamy distributions, along with the corresponding estimation errors and fit statistics.

Parameter/ Statistic	Series A			Series B			Series C		
	GLU	BETA	KUM	GLU	BETA	KUM	GLU	BETA	KUM
α	0.6603	0.8939	0.5989	2.3400	1.9597	1.5018	1.2055	4.7587	1.3589
β	1.1541	1.9902	1.3840	0.8773	1.1883	1.0948	1.5639	4.6025	1.7445
MSEE	0.0118	0.0153	0.0215	5.54×10^{-3}	7.98×10^{-3}	0.0426	2.86×10^{-3}	3.15×10^{-3}	0.0573
AIC	-116.0	-69.18	-83.81	-310.9	-145.5	-65.37	-1423.5	-1419.7	-218.0
BIC	-110.0	-63.13	-77.76	-294.3	-128.9	-48.78	-1404.8	-1401.0	-199.3
KS (p-value)	0.0921 (0.5393)	0.0987 (0.4498)	0.1316 (0.1439)	0.0623 (0.1654)	0.0886 * (0.0285)	0.1495 ** (1.11×10^{-3})	0.0392 (0.1858)	0.0403 (0.1580)	0.2398 ** (0.00)

* $0.01 < p < 0.05$; ** $p < 0.01$.

5. Conclusion

A novel, the so-called GLU distribution is presented here, along with its basic stochastic properties, where, among others, the asymmetry of the proposed distribution was shown. In addition, a quantile-based procedure was performed to estimate the parameters of the GLU distribution. The consistency and AN properties of the Q-estimators obtained in this way were considered, as well as a Monte-Carlo study of their efficiency. In order to further verify the effectiveness of the proposed model, it was applied in fitting the distributions of three real data sets, where it was compared with the Beta and Kumaraswamy distributions, thus demonstrating its higher efficiency. It is worth pointing out that the application of the GLU distribution shown here to training datasets intended for machine learning can also be the basis for some further research. Namely, as is known, logistic regressions are often used in machine learning and data science in general. Therefore, one of the guidelines for some potential further research can be the application of unit distributions based on logistic maps in this area.

Author Contributions: Conceptualization, V.S. and M.J.; methodology, V.S. and M.J.; software, V.S., M.J. and B.P.; validation, V.S., M.J. and B.P.; formal analysis, V.S. and M.J.; data curation, V.S., M.J. and Z.L.; writing—original draft preparation, V.S., M.J. and Z.L.; writing—review and editing, B.P. and Z.L.; visualization, V.S. and B.P.; supervision, B.P. and Z.L.; project administration, B.P. and Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original data presented in the study are openly available on the “Kaggle” website: <https://www.kaggle.com>.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Afify, A.Z.; Nassar, M.; Kumar, D.; Cordeiro, G.M. A New Unit Distribution: Properties and Applications. *Electron. J. Appl. Stat.* **2022**, *15*, 460–484.
2. Nasiru, S.; Abubakari, A.G.; Chesneau, C. New Lifetime Distribution for Modeling Data on the Unit Interval: Properties, Applications and Quantile Regression. *Math. Comput. Appl.* **2022**, *27*, Article No. 105. <https://doi.org/10.3390/mca27060105>
3. Martínez-Flórez, G.; Azevedo-Farias, R.B.; Tovar-Falón, R. New Class of Unit-Power-Skew-Normal Distribution and Its Associated Regression Model for Bounded Responses. *Mathematics* **2022**, *10*, Article No. 3035. <https://doi.org/10.3390/math10173035>
4. Bakouch, H.S.; Hussain, T.; Tošić, M.; Stojanović, V.S.; Qarmalah, N., Unit Exponential Probability Distribution: Characterization and Applications in Environmental and Engineering Data Modeling, *Mathematics* **2023**, *11(19)*, Article No. 4207. <https://doi.org/10.3390/math11194207>
5. Korkmaz, M. Ç.; Korkmaz, Z. S. The Unit Log-log Distribution: A New Unit Distribution With Alternative Quantile Regression Modeling and Educational Measurements Applications. *J. Appl. Stat.* **2023**, *50(4)*, 889–908.

6. Nasiru, S.; Abubakari, A.G.; Chesneau, C. The Arctan Power Distribution: Properties, Quantile and Modal Regressions with Applications to Biomedical Data. *Math. Comput. Appl.* **2023**, *28*, Article No. 25. <https://doi.org/10.3390/mca28010025>
7. Shakhathreh, M.K.; Aljarrah, M.A. Bayesian Analysis of Unit Log-Logistic Distribution Using Non-Informative Priors. *Mathematics* **2023**, *11*, Article No. 4947. <https://doi.org/10.3390/math11244947>
8. Fayomi, A.; Hassan, A.S.; Baaqeel, H.; Almetwally, E.M. Bayesian Inference and Data Analysis of the Unit-Power Burr X Distribution. *Axioms* **2023**, *12*, Article No. 297. <https://doi.org/10.3390/axioms12030297>
9. Alomair, G.; Akdoğan, Y.; Bakouch, H.S.; Erbayram, T. On the Maximum Likelihood Estimators' Uniqueness and Existence for Two Unitary Distributions: Analytically and Graphically, with Application. *Symmetry* **2024**, *16*, Article No. 610. <https://doi.org/10.3390/sym16050610>
10. Condino, F.; Domma, F.; Unit Distributions: A General Framework, Some Special Cases, and the Regression Unit-Dagum Models. *Mathematics* **2023**, *11*(13), Article No. 2888. <https://doi.org/10.3390/math11132888>
11. Krishna, A.; Maya, R.; Chesneau, C.; Irshad, M.R. The Unit Teissier Distribution and Its Applications. *Math. Comput. Appl.* **2022**, *27*, Article No. 12. <https://doi.org/10.3390/mca27010012>
12. Ramadan, A.T.; Tolba, A.H.; El-Desouky, B.S. A Unit Half-Logistic Geometric Distribution and Its Application in Insurance. *Axioms* **2022**, *11*, Article No. 676. <https://doi.org/10.3390/axioms11120676>
13. Mustafa, Ç.; Korkmaz, Z.; Korkmaz, S. The Unit Log-log Distribution: A New Unit Distribution With Alternative Quantile Regression Modeling and Educational Measurements Applications. *J. Appl. Stat.* **2023**, *50*(4), 889–908, DOI: [10.1080/02664763.2021.2001442](https://doi.org/10.1080/02664763.2021.2001442).
14. Nasiru, S.; Chesneau, C.; Abubakari, A.G.; Angbing, I.D. Generalized Unit Half-Logistic Geometric Distribution: Properties and Regression with Applications to Insurance. *Analytics* **2023**, *2*, 438–462. <https://doi.org/10.3390/analytics2020025>
15. Biçer, C.; Bakouch, H.S.; Biçer, H.D.; Alomair, G.; Hussain, T.; Almohisen, A. Unit Maxwell-Boltzmann Distribution and Its Application to Concentrations PoGLUtant Data. *Axioms* **2024**, *13*(4), Article No. 226. <https://doi.org/10.3390/axioms13040226>
16. Nasiru, S.; Chesneau, C.; Ocloo, S.K., The Log-Cosine-Power Unit Distribution: A New Unit Distribution for Proportion Data Analysis. *Decis. Anal. J.* **2024**, *10*, Article No. 100397
17. Alsadat, N.; Taniş, C.; Sapkota, L.P.; Kumar, A.; Marzouk, W.; Gemeay, A.M. Inverse Unit Exponential Probability Distribution: Classical and Bayesian Inference With Applications. *AIP Advances* **2024**, *14*, Article No. 055108.
18. Stojanović, V.S.; Jovanović Spasojević, T.; Jovanović, M. Laplace-Logistic Unit Distribution with Application in Dynamic and Regression Analysis. *Mathematics* **2024**, *12*, Article No. 2282. <https://doi.org/10.3390/math12142282>
19. Stojanović, V.S.; Bojičić, R.; Pažun, B.; Langović, Z. Quasi-Lindley Unit Distribution: Properties and Applications in Stochastic Data Modeling. *U.P.B. Sci. Bull., Series A*, submitted manuscript.
20. Rudin, W. Real and complex analysis (3rd ed.), McGraw-Hill International Edition, New York, 1987.
21. Salinas, H.S.; Bakouch, H.S.; Almuhayfith, F.E.; Caimanque, W.E.; Barrios-Blanco, L.; Albalawi, O. Statistical Advancement of a Flexible Unitary Distribution and Its Applications. *Axioms* **2024**, *13*, Article No. 397. <https://doi.org/10.3390/axioms13060397>
22. Dudek, D.; Kuczmaszewska, A. Some Practical and Theoretical Issues Related to the Quantile Estimators, *Stat. Papers* **2024**, in press, <https://doi.org/10.1007/s00362-024-01543-3>.
23. Serfling, R.J. Approximation Theorems of Mathematical Statistics, 2nd edition; John Wiley & Sons: New York, NY, USA, 2002.
24. Ruckdeschel, P.; Kohl, M.; Stabla, T.; Camphausen, F. S4 Classes for Distributions. *R News*, **2006**, *6*:(2–6). Available online: <https://CRAN.R-project.org/doc/Rnews> (accessed on 01 August, 2024).
25. Gross, L. Tests for Normality. *R package version 1.0-2*. <http://CRAN.R-project.org/package=nortest>, 2013. <http://CRAN.R-project.org/package=nortest> (accessed on 1 August, 2024).
26. Gumbel, E.J. Statistical theory of extreme values and some practical applications, *Appl. Math. Ser.* **1954**, *33*. U.S. Department of Commerce, National Bureau of Standards.
27. Burke, E.J.; Perry, R.H.J.; Brown, S.J. An extreme value analysis of UK drought and projections of change in the future. *J. Hydrol.* **2010**, *388* (1–2), 131–143. <https://doi.org/10.1016/j.jhydrol.2010.04.035>

28. Pažun, B., Langović, Z. Contemporary Information System Development Methodologies in Tourism Organizations, *4th International Scientific Conference, Tourism in Function of Development of the Republic of Serbia, Spa Tourism in Serbia and Experiences of Other Countries*, 2019, Vrnjačka Banja, Serbia, 467-481. <https://www.tisc.rs/proceedings/index.php/hitmc/article/view/267>
29. Kaggle.com <https://www.kaggle.com> (accessed on 2 August, 2024).
30. Mirza, M.H. "telecom_churn_dataset" (dataset), 2023. <https://www.kaggle.com/datasets/mirzahasnine/telecom-churn-dataset> (accessed on 2 August, 2024).
31. Mnassri, B. "Telecom Churn Dataset" (dataset), 2023. <https://www.kaggle.com/datasets/mirzahasnine/telecom-churn-dataset> (accessed on 2 August, 2024).
32. Kachiashvili, K.J.; Melikdzhanjan, D.I. Estimators of the Parameters of Beta Distribution. *Sankhya B* 2019, 81, 350–373. <https://doi.org/10.1007/s13571-018-0157-2>
33. Dey, S.; Mazucheli, J.; Nadarajah, S. Kumaraswamy Distribution: Different Methods of Estimation. *Comp. Appl. Math.* 2018, 37, 2094–2111. <https://doi.org/10.1007/s40314-017-0441-1>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.